Cognitive advantage in bilingualism

Cognitive advantage in bilingualism: An example of publication bias?

Angela de Bruin[1], Barbara Treccani[2], and Sergio Della Sala[1, 3]

1) Human Cognitive Neuroscience, Psychology, University of Edinburgh, Edinburgh, UK
2) Department of History, Human Sciences and Education, University of Sassari, Sassari, Italy
3) Centre for Cognitive Ageing and Cognitive Epidemiology, Psychology, University of Edinburgh, Edinburgh, UK

Address correspondence to:

Angela de Bruin

Human Cognitive Neuroscience, Psychology

University of Edinburgh

7 George Square

Edinburgh EH8 9JZ - United Kingdom

Phone: +44 (0) 131 650 3426

Fax: +44 (0) 131 651 3230

Email: angela.debruin@ed.ac.uk

Key words: Bilingualism, Executive functions, Inhibition, Publication bias

Abstract

It is a widely held belief that bilinguals have an advantage over monolinguals in executive control tasks, but is this an accurate reflection of all studies? The idea of a bilingual advantage may result from a publication bias favouring studies with positive results over null or negative effects. We examined the publication percentages of conference abstracts with results supporting or challenging the bilingual advantage. Studies with results fully supporting the bilingual-advantage theory were most likely to be published, followed by studies with mixed results. Studies challenging the bilingual advantage were published the least. This discrepancy was not due to differences in sample size, tests used, or statistical power. A test for funnel plot asymmetry provided further evidence for the existence of a publication bias.

1. Introduction

Many studies have suggested that bilinguals have an advantage over monolinguals in executive control. Evidence for this claim has been obtained in studies with children (Bialystok & Martin, 2004), young adults (Costa, Hernández, & Sebastián-Gallés, 2008), and older adults (Gold, Kim, Johnson, Kryscio & Smith, 2013), using tasks showing smaller interference effects in bilinguals than monolinguals, including Simon (Bialystok, Craik, Klein, & Viswanathan, 2004), flanker (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009), and task-switching paradigms (Prior & MacWhinney, 2010).

The argument of bilingualism enhancing cognitive control is also extensively discussed in book chapters (e.g., Bialystok & Barac, 2013), special Issues (e.g., Bobb, Wodniecka, & Kroll, 2013; Kroll, Christoffels, & Bajo, 2013), and conferences (e.g., 'Bilingualism and cognitive control', http://www.blcc2013.langusta.edu.pl/). Based on these studies, the media have often presented a picture of a strong bilingual advantage, represented by headings such as 'Bilingual brains are more healthy' (Fox, 2011) or 'Why bilinguals are smarter' (Bhattacharjee, 2012), suggesting that that the idea is consolidated and accepted as common wisdom. Despite this ongoing belief, not all studies have found an advantage of bilinguals over monolinguals. Some of these studies have been published (e.g., Paap & Greenberg, 2013; Gathercole et al., 2014), but we suspected that many other studies have remained unpublished.

We ourselves are guilty. We contributed to create the accepted wisdom of a cognitive advantage in bilinguals by publishing a study reporting an effect of bilingualism in a spatial negative priming task (Treccani, Argyri, Sorace, & Della Sala, 2009). This effect, supporting the theories of enhanced inhibitory control in bilinguals, was obtained in one experiment. Three other tasks (Simon, colour negative priming, and spatial cueing tasks), however, were administered at the same time and to the same participants and did not show any differences between bilinguals and monolinguals. The only experiment that we submitted for publication was the one showing an effect of bilingualism. Similarly, another study from our research group (using the same spatial negative priming paradigm that was successful in Treccani et al., 2009) failed to replicate the observed effect of bilingualism. Due to the same file drawer bias (cf., Spellman, 2012), this study was not submitted either.

We then wondered if the claim that bilinguals have a cognitive advantage is a correct reflection of all research in this field. Recently, Paap (2014) has raised the concern that the literature on bilingualism and executive control might be affected by a confirmation bias to report positive results only. In fact, the file drawing problem is not the only bias that marks scientific literature: The well-known 'publication bias' is another obvious one (cf., Chambers, 2013; Cumming, 2013; Fanelli, 2010; Easterbrook, Gopalan, Berlin, & Matthews, 1991; Francis, 2012). To investigate whether and to what extent studies showing a bilingual advantage are more likely to be published than data challenging the bilingual-advantage hypothesis, we compared the publication rates of conference abstracts. We classified conference abstracts on the basis of their outcomes and assessed which abstracts were subsequently published in a journal.

2. Methods

We searched for conference abstracts on bilingualism and executive control in 169 conferences (31 different national and international meetings) organised between 1999 and 2012. The topics of these conferences included bilingualism, psycholinguistics, cognitive neuroscience, psychology, and psychiatry (see Table 1 for an overview of all conferences).

Table 1

Overview of the different conferences at which the analysed abstracts were presented.

Conference	Year
American Aging Society	2005 - 2011
Architectures and Mechanisms for Language Processing (AMLaP)	2002 – 2004; 2007 - 2011
Association for Psychological Science (APS)	2003 - 2012
Boston University Conference on Language Development (BUCLD)	2008 - 2012
Canadian Society for Brain, Behaviour, and Cognitive Science (CSBBCS)	2004 - 2012
Cognitive Science Society (CogSci)	2003 - 2012
Cognitive Neuroscience Society (CNS)	2003 - 2012
CUNY Sentence Processing Conference	2006 - 2012
European Brain and Behavior Society (EBBS)	2003 - 2009
European Congres of Psychology (ECP)	2009
European Federation of Neurological Societies (EFNS)	2005 - 2011
European Society for Cognitive and Affective Neuroscience (ESCAN)	2012
European Society for Cognitive Psychology (ESCOP)	2007; 2009; 2011
FENS Forum of Neuroscience	2002; 2004; 2006; 2008; 2010; 2012
International Association for the Study on Child Language (IASCL)	2005; 2008; 2011
International Conference on Cognitive Neuroscience (ICON)	2011
International Conference on Models of Interaction in Bilinguals (ICMIB)	2009
International Neuropsychological Society (INS)	2003 - 2010
International Symposium on Bilingualism (ISB)	2003; 2007; 2009; 2011
International Symposium of Psycholinguistics (ISP)	2011
Midwestern Psychological Association (MPA)	2004; 2005; 2006; 2010; 2011; 2012
Neurobilingualism	2009
Neurobiology of Language (NBL)	2009 – 2012
Nordic Conference on Bilingualism	2009; 2012
Psychonomics	1999 – 2012
Society for Neuroscience (SNF)	2000 – 2012
Society for Psychophysiological Research (SPR)	2004-2012
Society for Research in Child Development (SRCD)	2005; 2007; 2009; 2011
Workshop on Bilingualism	2005 – 2008; 2011
Workshop on Neurobilingualism	2010
Bilingual & Multilingual Interaction	2012

We identified 128 abstracts (presented at 52 different conferences) that focussed on bilingualism and executive control. We included all abstracts that investigated the relationship between bilingualism and executive control in any age group, either with non-linguistic control tasks (116 abstracts; both standard executive control tasks, e.g., the Simon task, or tasks with a clear executive control component, e.g., working memory updating tasks) or with linguistic control tasks (12 abstracts, e.g., homograph interference task). We included executive control tasks with linguistic stimuli to get a complete overview of the publication bias in the general field of bilingualism and executive control. We did not include conference abstracts looking at effects of bilingualism in lexical tasks without a clear executive control component (e.g., word learning or picture naming tasks). A total of 24 conference abstracts could not be classified because the abstract did not contain enough information about the results (15 abstracts), the study was lacking a (monolingual) control group (8 abstracts), or because the abstract was a review of previous studies (1 abstract). Two authors classified independently the remaining 104 abstracts according to their reported results. Any disagreement, which occurred in 11 cases, was resolved by discussion.

2.1. Classification

We classified the abstracts into four categories (see Supplementary Materials for an overview of all abstracts and their classifications):

1. Studies only reporting data that support the bilingual advantage (‘yes’ studies).
2. Studies reporting mixed data that, on the whole, support the bilingual-advantage hypothesis (‘mixed-yes’ studies). These studies do not report a bilingual advantage in all tasks/analyses, but their results are still compatible with the prevalent idea of bilinguals showing enhanced abilities in executive control (they report no bilingual advantage in experimental conditions where an effect of bilingualism was not expected). This includes studies that provide neuroimaging or electrophysiological evidence for a monolingual-bilingual difference in executive control consistent with the idea of more efficient executive functions in bilinguals and studies that show bilingual advantages (a) for high executive control conditions (e.g., in flanker tasks involving strong interference effects), but not for low executive control conditions (e.g., in flanker tasks involving weaker interference effects) (5 studies); (b) for executive control tasks where a bilingual advantage was expected (e.g., domain-

general control tasks such as Simon tasks), but not in other tasks where no bilingual advantage was expected (tasks in which performance depends on expertise in a particular field such as music, or tasks tapping executive functions that are not directly related to the ability of controlling two or more antagonist cognitive networks such as the two languages in a bilingual, e.g., the impulse-delay task; 6 studies); (c) for high proficiency bilinguals, but not low proficiency bilinguals (1 study) or for switching balanced bilinguals, but not for non-switching balanced bilinguals (1 study); (d) for unimodal, but not bimodal bilinguals (i.e., people proficient in one spoken language and one sign language; 1 study).

3. Studies reporting mixed data that partly challenge the bilingual advantage ('mixed-no' studies). These studies report some results that support the bilingual advantage idea, but also report experiments where a bilingual advantage was expected, but not found or data indicating that the bilingual advantage in some tasks could have other explanations than the mere knowledge of two languages and the ability to master them. This includes studies that show (a) a bilingual advantage in some executive control tasks, but not in other (parts of the) tasks where an effect of language group was expected too (20 studies); (b) a bilingual advantage for certain language groups, but not others (5 studies); (c) some (inconsistent) effects of language group in neuroimaging or electrophysiological data, but no bilingual advantage in behavioural data (reaction times; RTs) (6 studies); (d) a bilingual advantage for some age groups but not others (1 study); (e) a bilingual advantage that could be explained by other factors such as the socio-economic status of participants (1 study).
4. Studies reporting results that fully challenge the bilingual advantage ('no' studies): Studies that do not show any difference between monolinguals and bilinguals and studies that demonstrate a bilingual disadvantage.

We have based our classification on the results and conclusions reported in the *conference abstracts*. In some cases, the study described in the abstract ended up being published in a scientific journal and the conclusions drawn by the author in the abstract do not match those drawn in the published paper. For example, the abstract by Luk, Anderson, Craik, Grady, and Bialystok (2009; see Supplementary Materials) does not discuss the absence of a bilingualism effect on RTs, but only focuses on the bilingual 'advantage'

observed in fMRI data. Based on this abstract, we have classified their study as belonging to group 1 ('yes' studies). These authors also describe this study in a published paper (Luk, Anderson, Craik, Grady, & Bialystok, 2010), in which they mention the absence of an RT effect. Based on this paper, a classification in group 3 ('mixed-no') would have been more appropriate. To avoid differences between published and unpublished studies, however, we based our categorisation on conference abstracts only.

After classifying the abstract, we identified whether the results presented in the conference abstract had been published in a journal. We classified a paper as being published if, on the 20th of February 2014, it had been accepted for publication by an international scientific journal. Either papers that had already appeared in a journal issue or in-press papers were classified as published. We did not include book chapters or published conference proceedings. We also classified an abstract as published if the results were part of a paper with additional experiments or participants. If two conference abstracts from the same research group, which reported different studies (e.g., Paap et al., 2010, and Paap et al., 2012), were later combined to form one journal publication, we classified both abstracts as being published. However, when two abstracts presented at different conferences reported exactly the same study findings, only the first conference presentation was deemed suitable for inclusion.

We also identified three factors that could potentially confound the results: year of conference presentation, number of participants per language group, and number of executive-control tasks administered in the study. We included the number of participants per group rather than the total number of participants as some studies included many different groups (e.g., different language combinations) or multiple tasks, thus leading to very high numbers of participants.

2.2. Meta-analysis

We performed a meta-analysis of the published studies, which provided suitable data, and assessed the presence of a publication bias by means of a funnel plot. Of the 50 identified published papers from our conference abstracts-assessment, we included 41 papers in our meta-analysis. We contacted the authors if the paper did not provide the required descriptive statistics. Nine studies could not be included in the analysis, because we could not obtain the descriptive statistics, the paper focussed on neuroimaging data only, or the

author did not allow inclusion of their study's results in the analysis. We included all behavioural executive control tasks described in the papers, but did not include neuroimaging data, and only analysed bilingual-monolingual differences on the critical dependent variables (e.g., if the paper focussed on RTs, we only included RT, but not accuracy results). For tasks that reported overall RTs as well as conflict effects (e.g., Simon or flanker task), we only included the conflict effects. If the study compared multiple bilingual or monolingual groups, we analysed those groups separately. In total, our analysis contained 176 comparisons. We used MetaXL 2.0 software and the metafor software package in R (Viechtbauer, 2010) for our statistical analysis.

3. Results

3.1. Conference abstracts

Of the 104 abstracts included in our analysis, 40 abstracts (38%) found a bilingual advantage or results supporting the bilingual-advantage theories. A total of fourteen studies found mixed results supporting the bilingual advantage theories (13%). Thirty-three studies showed mixed results partly challenging the bilingual advantage theories (32%). Seventeen studies (16%) found no differences between monolinguals or bilinguals (13 studies) or a monolingual advantage (4 studies). In total, 52 studies were published in 50 papers (50% of all conference abstracts). Sixty-eight per cent of the 'yes' studies were published, compared to 50% of the mixed-yes studies, 39% of the mixed-no studies, and 29% of the no studies. On the whole, 63% of the studies supporting the bilingual advantage were published compared to only 36% of the challenging studies (see Table 2 and Figure 1).

Table 2

Overview of the number of abstracts supporting and (either fully or partially) challenging the bilingual advantage that were presented at international conferences (1999-2012). Number and percentage of studies that ended up in publication are also presented. See paragraph 2.1 for a description of the four result types ('yes', 'mixed-yes', 'mixed-no', 'no').

Result type	N abstracts	N published	% published
1. Bilingual advantage ('yes')	40	27	68
2. Mixed data supporting bilingual advantage theory ('mixed-yes')	14	7	50
3. Mixed data (partly) challenging bilingual advantage theory ('mixed-no')	33	13	39
4. No bilingual advantage ('no')	17	5	29
- *Bilingual disadvantage*	*4*	*2*	*50*
- *No differences between monolinguals and bilinguals*	*13*	*3*	*23*
Results supporting the bilingual advantage (1+2)	54	34	63
Results challenging the bilingual advantage (3+4)	50	18	36

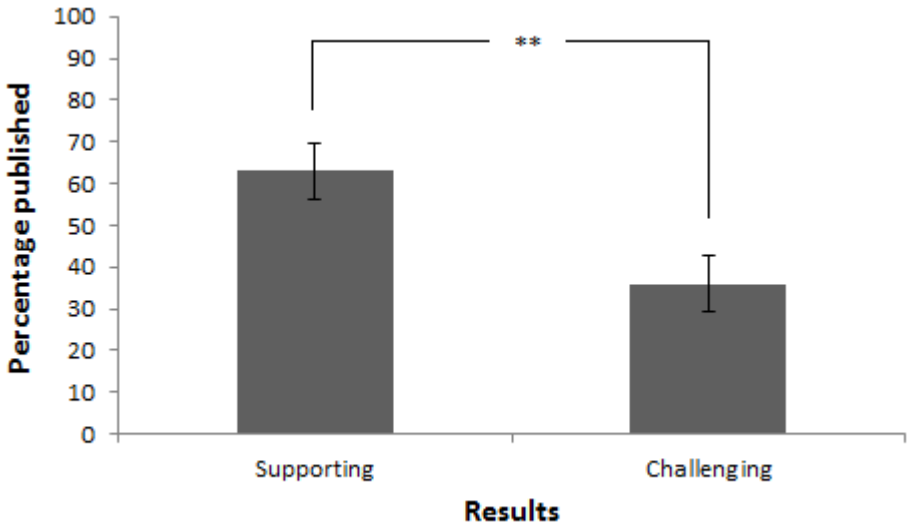

Figure 1. Percentage of conference papers supporting or challenging the bilingual advantage that were published in an international scientific journal. Error bars show +/- standard error of the mean. **: $p < .01$.

Using a binary logistic regression analysis, we found a significant difference between the publication outcomes (published or unpublished) of bilingual-advantage challenging and supporting abstracts, Wald χ^2 =7.36, df = 1, p = .007, η_p^2 = .073. When we included all four result types, the analysis still showed a significant effect of result type on publication, Wald χ^2 =8.86, df = 3, p = .031, η_p^2 = .089.

Using independent t-tests, we found no significant differences between the bilingual-advantage supporting and challenging abstractsin terms of year of conference presentation and number of participants per group (see Table 3). Challenging abstracts, however, reported more executive control tasks than supporting abstracts. Not all abstracts included information on the number of executive control tasks and number of participants per group. Among the abstracts supporting the bilingual advantage, 9 studies did not report information on the number of participants and 3 abstracts lacked detail on the number of tasks. Among the abstracts challenging the bilingual advantage, 15 studies did not include information on the number of participants and 3 abstracts on the number of tasks. These analyses, therefore, include the majority of studies, but not all studies, and the results should be interpreted with caution.

Table 3

Means (and standard deviations) of the year of the conference at which the analysed studies (either supporting or challenging the bilingual advantage) were presented, number of participants per language group, and number of tasks administered in the study.

	Results		Significant difference?
	Supporting	Challenging	
Year of conference presentation	2008.9 (1.97)	2009.2 (2.76)	No (t = .50, p = .620)
Number of participants per group	31.1 (23.76)	28.3 (16.21)	No (t = .60, p = .554)
Number of tasks	1.6 (1.29)	2.2 (1.25)	Yes (t = 2.35, p = .021)

3.2. Meta-analysis

Our meta-analysis of the published studies showed an effect of bilingualism with an average standardised mean difference of 0.30 (95% CI 0.23 to 0.37, z = 8.21, p < .0001; see Supplementary Figure 1 for the forest plot). Importantly, the funnel plot (i.e., a scatter plot in which the bilingualism effects estimated from individual studies - standardised bilingual-monolingual mean differences - on the horizontal axis are plotted against a measure of study precision - standard errors - on the vertical axis) shows a clear asymmetry (see Figure 2). Studies with larger standard errors showed a larger standardised mean difference than studies with smaller standard errors. In the absence of a publication bias, the funnel plot should have been symmetrical with studies with larger standard errors resulting in a similar amount of relatively high and low standardised mean differences. Studies with larger standard errors should then scatter widely at the bottom of the graph (cf., Sterne, Becker, & Egger, 2005). Instead we observe that less precise studies (with larger standard errors) more often show large effects than small effects, which suggests that studies with small effect sizes might not have been published. The observed asymmetry in the funnel plot was

further supported by Egger's linear regression test, which showed a significant asymmetry (z = 4.80, p < .0001).

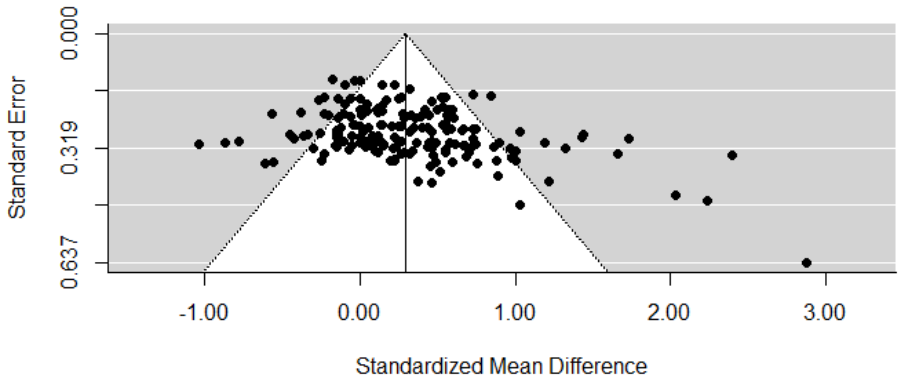

Figure 2. Funnel plot of the meta-analysis of published papers.

We also calculated the power of the analysed null effect studies to detect the various effect sizes reported in published positive papers. We calculated the power for studies concerning the Simon effect, flanker effect, and task switching costs. On the whole, null effect studies had a medium-to-high probability of detecting the positive effects reported by published studies using the same tasks. For example, in the Simon task used by Bialystok et al. (2004), the bilingualism effect size (Cohen's *d*) ranged from 1.08 to 2.99. Using G*Power 3.1.8. (Faul, Erdfelder, Lang, & Buchner, 2007) with an alpha level of .05 (two tails), we calculated the probability to detect this effect, that is, the statistical power (1-β) of the null effect abstracts included in our overview that used a Simon task and provided sufficient information (i.e., they provided the number of participants per group; 12 studies; see Supplementary Materials for full references). All studies using Simon tasks analysed here had very high probability to detect such a large effect (average of .87 to detect *d* of 1.08, and .99 to detect *d* of 2.99). The effect sizes for Simon effects reported by two other published positive studies (i.e., Bialystok, Craik, and Luk, 2008, and Salvatierra and Rosselli , 2011) were smaller (.59 and .69 respectively) and null effect studies had a medium probability to detect them (average of .52 and .66 respectively). The same procedure was used for flanker and task-switching studies. In the flanker task used by Costa et al. (2009) the effect size of bilingualism was .61. Null effect flanker-task studies (8 studies)

on average had a medium probability (.62) of detecting this effect. In a task-switching paradigm, Gold et al. (2013) found a bilingualism effect size of .68 and null effect task-switching studies (3 studies) on average had a high probability of detecting it (.94).

Finally, we calculated the power of both supporting and challenging abstracts to detect the effect size found in our meta-analysis (.30). Eighty studies (45 supporting and 35 challenging studies, see supplementary materials for references) provided sufficient information to be included in the analysis. For studies classified as supporting the bilingual advantage, the power to detect an effect size of .3 was .19. For studies classified as challenging the bilingual advantage, this power was .17. Both types of studies thus had a comparable, but low probability to detect the effect size observed in the meta-analysis.

4. Discussion

We have analysed conference abstracts presented between 1999 and 2012 on the topic of bilingualism and executive control. Conference abstracts were classified on the basis of their outcome. We observed an effect of result type on publication: Studies were published relatively often (68%) if the data demonstrated a bilingual advantage. In contrast, only 29% of the studies that showed no effect of bilingualism or even a bilingual disadvantage were published. Publication chances of studies with mixed results were in-between these two groups, with studies partly supporting the bilingual advantage being more likely to be published than partly challenging studies. The asymmetrical funnel plot of published studies also hinted at a publication bias.

This difference in publication percentage based on the outcomes of the study could be the result of a bias during several steps of the publication process: Authors, reviewers, and editors can decide to only submit or accept studies with positive results.

In the first step of the publication process, the file drawer problem could play an important role in the observed publication bias. Authors could decide not to publish studies with null or mixed results or they can choose to submit their results only partially, for example by leaving out tasks that did not show an effect of bilingualism. The paper by Treccani et al. (2009) is an example of file drawing as it excluded the experiments that did not show an effect of bilingualism.

On the next level, reviewers and/or editors might reject a submitted paper reporting null, negative, or mixed results more often than studies finding positive effects. This rejection is often based on the argument that null effects are difficult to interpret, the result of poor stimulus design, or the result of a Type-II error (Ferguson & Heene, 2012). Mahoney (1977) asked journal reviewers to referee studies reporting positive, negative, mixed, or null results with identical methodological procedures. Although the methodology was the same, reviewers scored the positive papers as methodologically better than the negative or mixed results papers. For papers with positive results, reviewers usually recommended accepting with moderate revisions. For papers with negative results, however, their usual recommendation was major revision or rejection. Papers with mixed results were mostly rejected.

Unfortunately, we cannot determine whether studies were not submitted to a journal or rather rejected after submission. We did ask all authors to take part in a short survey concerning journal submission, but only a small proportion responded. This lack of responsiveness, particularly from very productive research groups, complicates the interpretation of our findings.

In our overview of conference abstracts, almost half of the abstracts (48%) partially challenged the existence of a bilingual advantage by reporting tasks with no effect of bilingualism. We should be careful interpreting null effects. It is worth noticing, however, that most studies analysed here (showing positive, mixed or null effects) have used the same tasks (e.g., Simon, task-switching, or flanker task). Our analyses furthermore showed that studies supporting the bilingual-advantage hypothesis and studies challenging this hypothesis did not differ significantly in terms of number of participants per group. The two types of studies also had a similar power to detect the bilingualism effect size found in the meta-analysis. Interestingly, we did observe a difference between the different abstract types in the number of reported tasks. Abstracts supporting the bilingual advantage reported fewer tasks than abstracts with mixed results or abstracts challenging the bilingual advantage. It is difficult to interpret this difference, as it might reflect a difference in the number of tasks that were reported rather than a difference in the number of tasks that were actually used. Although this is speculative, the difference in number of tasks between these studies could be the result of some of the 'supporting' studies leaving out data that suggested otherwise (cf., John, Loewenstein, & Prelec, 2012). Alternatively, a significant

effect could be most likely to occur if only one test is used, whereas more tests might also yield non-significant or negative results. Researchers could submit a paper after one successful task without trying to replicate this effect, even if the positive outcome could be the result of a Type-I error (cf., Pashler & Harris, 2012).

Only a few of the analysed studies (4 of the 104 abstracts) found a bilingual *disadvantage.* A lack of this kind of abstracts could result from file drawing on the level of conference submission already. Indeed, the finding of a bilingual disadvantage can hardly be interpreted as indicating better executive control abilities in monolinguals. The only reasonable conclusions would be that, in the tested domain, there is no bilingual advantage and a Type-I error occurred. Authors then might not submit their negative results to a conference. In this respect, it is worth noticing that file drawing occurring at conference-submission level might have obscured the existence of differences in publication rates even larger than those we found: Our results might only be the 'tip of the iceberg'.

The small percentage of bilingual disadvantage studies, however, could also suggest that the cognitive bilingual advantage is genuine, albeit smaller and less stable than often presented in the literature. In fact, the existence of a publication bias does not necessarily imply that bilingualism does not have any effect on executive functions. The presence of a possible publication bias may explain why the magnitudes of many reported positive effects appear to decrease over time (i.e., their size declines as studies exploring them are repeated), even when the effects have been shown to be reliable and are still widely-accepted (cf., Schooler, 2011; see also Lehrer, 2010). Yet, the presence of a possible publication bias can help to interpret the exaggerated outcomes often reported in the initial phase of research.

Our overview shows that there is a distorted image of the actual study outcomes on bilingualism, with researchers (and media) believing that the positive effect of bilingualism on non-linguistic cognitive processes is strong and unchallenged. Recently, however, several studies (e.g., Paap, 2014; Paap & Liu, 2014) have criticised the findings in the existing literature. Their criticisms focus especially on the impossibility to assign randomly the independent variable (i.e., language group), and on the differences between bilingual and monolingual groups on background variables such as the socio-economic or immigration status. In light of these issues, it is especially important to avoid publishing positive studies only.

A potential publication bias also poses a problem for meta-analyses. On the basis of an estimation of the number of *possible* unpublished null-effect studies, Adesope, Lavin, Thompson, & Ungerleider (2010) concluded that it was unlikely that their meta-analysis on bilingualism cognitive effects could be threatened by a publication bias. Conversely, our overview shows the number of *actually conducted* unpublished null-effect studies and suggests that the results of a meta-analysis can be in fact affected by such a bias. Hilchey and Klein (2011) reviewed published studies that specifically address the issue of bilingualism and executive control. Although this review rightfully criticised some of the current theories, it is still necessarily based on published work only. Similarly, our meta-analysis did show an effect of bilingualism, but for the aim of the funnel plot, we only included published studies. The bilingual advantage found in this meta-analysis would be smaller if the unpublished abstracts (with more null and negative effects) were included too.

If we want to get a better understanding of the actual effect of bilingualism and the boundaries of this effect, publication chances should not depend on the direction of the study results. Studies with mixed results, for example, are especially valuable because they can identify the circumstances under which a bilingualism effect may and may not occur, but, as shown by our analysis, they are published less often than studies that report data in favour of the bilingual advantage. Furthermore, studies showing no effects of bilingualism are often unfairly criticised. Recently, Kroll and Bialystok (2013) claimed that 'The considerable literature that reports group differences between monolingual and bilingual participants is greatly more informative than the attempted replications that fail to find significance.' (p. 502) and '… unless all conditions have been accounted for and all other explanations have been exhausted, it is misleading to call into question the reliability of the phenomena themselves.' (p. 503). While we agree that bilingualism should be conceived, a priori, as a positive and desirable achievement, we are also convinced that educational and political debates addressing the relevance of bilingualism should not be promoted by ignoring null or negative results. Instead of selecting exclusively those tasks and results that support current theories, investigators should make an attempt to include all conducted tasks and reported findings. On the other hand, reviewers and editors should be more open to studies that challenge the existing theories, especially when these are not yet fully established. We should share all data and let them speak for themselves, also and especially in issues like bilingualism for its enormous societal relevance and implications.

Author contributions. All authors contributed to the study concept and design. Data collection and analysis were performed by A. de Bruin. Abstract classification was performed by A. de Bruin and B. Treccani. A. de Bruin drafted the manuscript and B. Treccani and S. Della Sala provided critical revisions. All authors approved the final version of the manuscript for submission.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*(2), 207-245.

Bhattacharjee, Y. (2012, March 17). Why bilinguals are smarter. *The New York Times.* Retrieved from http://www.nytimes.com/2012/03/18/opinion/sunday/the-benefits-of-bilingualism.html

Bialystok, E., & Barac, R. (2013). Cognitive effects. In F. Grosjean, & P. Li (Eds.). *The Psycholinguistics of Bilingualism.* (pp. 192-213). Chichester, UK: Wiley-Blackwell.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, *19*(2), 290-303.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 859-873.

Bialystok, E., & Martin, M.M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7,* 325-339.

Bobb, S., Wodniecka, Z., & Kroll, J. (Eds). (2013). What bilinguals tell us about cognitive control [Special issue]. *Journal of Cognitive Psychology, 25*(5).

Chambers, C. D. (2013). Registered Reports: A New Publishing Initiative at Cortex. *Cortex, 49,* 609-610.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135-149.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59-86.

Cumming, G. (2013). The New Statistics: Why and How. *Psychological Science.* DOI: 10.1177/0956797613504966

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867-872.

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS One*, *5*(4), 1-10.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*(6), 555-561.

Fox, K. (2011, June 12). Ellen Bialystok: Bilingual brains are more healthy. *The Guardian.* Retrieved from http://www.theguardian.com/technology/2011/jun/12/ellen-bialystok-bilingual-brains-more-healthy

Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin & Review*, *19*(2), 151-156.

Gathercole, V. C., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas-guasch, N., Roberts, E. J., Hughes, E. K., & Jones, L. (2014). Does language dominance affect cognitive performance in Bilinguals? Lifespan evidence from preschoolers through older adults on Card Sorting, Simon, and Metalinguistic Tasks. *Frontiers in Psychology*, *5*(11), 1-14.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience*, *33*(2), 387-396.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review*, *18*(4), 625-658.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524-532.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.

Kroll, J. F., Christoffels, I., & Bajo, T. (Eds.). (2013). Bilingualism and cognitive control [Special issue]. *Frontiers in Psychology, 2.*

Lehrer, J. (2010, December 13). The truth wears off. *The New Yorker.*

Retrieved from http://www.newyorker.com/reporting/2010/12/13/101213fa_fact_lehrer?currentPage=all

Luk, G., Anderson, J. A., Craik, F. I., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and cognition*, *74*(3), 347-357.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive therapy and research*, *1*(2), 161-175.

Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology*. Advance online publication. DOI: 10.1080/20445911.2014.891597

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive psychology*, *66*(2), 232-258.

Paap, K. R., & Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: The role of confirmation bias in testing for bilingual advantages. *Journal of Neurolinguistics*, *27*(1), 50-74.

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, *7*(6), 531-536.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(02), 253-262.

Salvatierra, J. L., & Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism*, *15*(1), 26-37.

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature, 470,* 437.

Spellman, B. A. (2012). Introduction to the Special Section: Data, Data, Everywhere... Especially in My File Drawer. *Perspectives on Psychological Science*, *7*(1), 58-59.

Sterne, J. A. C., Becker, B. J., Egger, M. (2005). The Funnel Plot. In H. Rothstein, A. Sutton, Borenstein, M. (Eds.), *Publication bias in meta-analysis: prevention, assessment, and adjustments* (75-98). West Sussex: John Wiley & Sons Ltd.

Treccani, B., Argyri, E., Sorace, A., & Della Sala, S. (2009). Spatial negative priming in bilingualism. *Psychonomic bulletin & review*, *16*(2), 320-327.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package.

Journal of Statistical Software, 36(3), 1-48.